



# COMSATS University Islamabad, Abbottabad Campus

## Department of Computer Science

### Examination – Midterm (Fall 2022)

Subject: Machine Learning (CSC354)

Total Time Allowed: 75 mins

Name: \_\_\_\_\_

Date: 14 November 2022

Instructor: Sara Shafique

Max Marks: 35

Registration # \_\_\_\_\_

Show all work clearly and legibly. Remember, you are being tested! So even if an answer is obvious to you, please show all the justifications by clearly showing the calculations, or explaining why a calculation is skipped.

## Section B

Answer all questions.

- Explain the difference between traditional programming and machine learning. [2 Marks]
  - Compare Classification with Regression using an example. [2 Marks]
- Determine type 1, type 2 errors and accuracy for the given confusion matrix. [3 Marks]

|               |              | Predicted Values |            |           |           |
|---------------|--------------|------------------|------------|-----------|-----------|
|               |              | Setosa           | Versicolor | Virginica | Row Total |
| Actual Values | Setosa       | 5                | 0          | 0         | 5         |
|               | Versicolor   | 0                | 3          | 1         | 4         |
|               | Virginica    | 0                | 1          | 5         | 6         |
|               | Column Total | 5                | 4          | 6         | 15        |

- What is the precision and recall tradeoff? [2 Marks]
  - Calculate precision and sensitivity for a given paradox. [2 Marks]

|        |          | Predicted |                   |          |
|--------|----------|-----------|-------------------|----------|
|        |          | Negative  | Positive          |          |
| Actual | Negative | 8<br>7    | 3<br>2            | TN<br>FP |
|        | Positive | 5<br>FN   | 5<br>5<br>5<br>TP |          |

4. a) Explain correlation types with plots.  
b) Normalize the given data.

[2 Marks]

[2 Marks]

|   | Student | CGPA | Salary '000 |
|---|---------|------|-------------|
| 0 | 1       | 3.0  | 60          |
| 1 | 2       | 3.0  | 40          |
| 2 | 3       | 4.0  | 40          |
| 3 | 4       | 4.5  | 50          |
| 4 | 5       | 4.2  | 52          |

## Section C

5. a) Distinguish between supervised learning and Reinforcement learning. Illustrate with an example. [5 Marks]

- b) Discuss any four examples of machine learning applications. [5 Marks]

6. a) Explain the procedure for computing principal components of the data. [5 Marks]

- b) Let's consider a hypothetical case for feature selection where we test the effectiveness of a drug for a certain medical condition. Suppose we have 105 patients under study and 50 of them were treated with the drug. The remaining 55 patients were kept as control samples. The health condition of all patients was checked after a week. The following table shows if their condition improved or not. [5 Marks]

|             | Responded | Not Responded | Total |
|-------------|-----------|---------------|-------|
| Treated     | 35        | 15            | 50    |
| Not Treated | 26        | 29            | 55    |
| Total       | 61        | 44            | 105   |